Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: https://orca.cardiff.ac.uk/id/eprint/111414/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Cowley, Laura E., Maguire, Sabine, Farewell, Daniel M. ORCID: https://orcid.org/0000-0002-8871-1653, Quinn-Scoggins, Harriet D. ORCID: https://orcid.org/0000-0002-6136-070X, Flynn, Matthew O. and Kemp, Alison M. ORCID: https://orcid.org/0000-0002-1359-7948 2018. Acceptability of the Predicting Abusive Head Trauma (PredAHT) clinical prediction tool: A qualitative study with child protection professionals. Child Abuse and Neglect 81 , pp. 192-205. 10.1016/j.chiabu.2018.04.022 file

Publishers page: http://dx.doi.org/10.1016/j.chiabu.2018.04.022 <http://dx.doi.org/10.1016/j.chiabu.2018.04.022>

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

information services
gwasanaethau gwybodaeth

Abstract

The validated Predicting Abusive Head Trauma (PredAHT) tool estimates the probability of abusive head trauma (AHT) based on combinations of six clinical features: head/neck bruising; apnea; seizures; rib/long-bone fractures; retinal hemorrhages. We aimed to determine the acceptability of PredAHT to child protection professionals. We conducted qualitative semi-structured interviews with 56 participants: clinicians (25), child protection social workers (10), legal practitioners (9, including 4 judges), police officers (8), and pathologists (4), purposively sampled across southwest United Kingdom. Interviews were recorded, transcribed and imported into NVivo for thematic analysis (38% double-coded). We explored participants' evaluations of PredAHT, their opinions about the optimal way to present the calculated probabilities, and their interpretation of probabilities in the context of suspected AHT. Clinicians**,** child protection social workers and police thought PredAHT would be beneficial as an objective adjunct to their professional judgment, to give them greater confidence in their decisions. Lawyers and pathologists appreciated its value for prompting multidisciplinary investigations, but were uncertain of its usefulness in court. Perceived disadvantages included: possible over-reliance and false reassurance from a low score. Interpretations regarding which percentages equate to 'low', 'medium' or 'high' likelihood of AHT varied; participants preferred a precise % probability over these general terms. Participants would use PredAHT with provisos: if they received multi-agency training to define accepted risk thresholds for consistent interpretation; with knowledge of its development; if it was accepted by colleagues. PredAHT may therefore increase professionals' confidence in their decision-making when investigating suspected AHT, but may be of less value in court.

Keywords: Child physical abuse*,* Abusive head trauma, Clinical prediction tool, Qualitative research, Child protection

Introduction

Abusive head trauma (AHT) has severe consequences for young children, and may be missed in the clinical setting (Jenny, Hymel, Ritzen, Reinert, & Hay, 1999; Letson et al., 2016). It is vital that AHT is accurately identified, to prevent further abuse from occurring and reduce the risk of falsely accusing innocent care providers. The evaluation and investigation of suspected AHT requires a multidisciplinary team approach whereby pediatricians collaborate with clinicians from other specialities, child protection social workers (CPSWs), law enforcement, and the justice system (Canadian Paediatric Society, 2007). Recent guidelines for the evaluation of suspected physical abuse recommend that medical records include a clear opinion about the likelihood of physical abuse and should elucidate specific levels of concern to aid police and CPSWs' investigations (Christian & Committee on Child Abuse & Neglect, 2015), so that they can gauge the degree of certainty of AHT in each case (Canadian Paediatric Society, 2007).

Following a systematic review (Maguire et al., 2009), derivation (Maguire, Kemp, Lumb, & Farewell, 2011) and external validation studies (Cowley, Morris, Maguire, Farewell, & Kemp, 2015), we developed the Predicting Abusive Head Trauma (PredAHT) clinical prediction tool (CPT). The derivation study used multivariable logistic regression and provided predicted probabilities of AHT for children (< 3 years old) with intracranial injury and each of 64 possible combinations of the presence or absence of six clinical features, detailed in Table 1 (S. Maguire et al., 2011). PredAHT performed with a sensitivity of 72.3% and a specificity of 85.7% in the validation study, using a cut-off probability value of 50% (Cowley et al., 2015). However, in the clinical setting some investigations may not be undertaken to identify key features e.g. X Rays or ophthalmology for fractures or retinal hemorrhages respectively. Using data from the derivation study (Maguire et al., 2011) we therefore estimated the probability of AHT when one or more features were unknown using

multiple imputation by chained equations (van Buuren & Groothuis-Oudshoorn, 2011). PredAHT thus estimates the probability of AHT for all 729 permutations of the six clinical features depending on whether each is present, absent *or unknown*. A computerised version was built using Shiny, a web application framework for the R language and environment for statistical computing (Chang, Cheng, Allaire, Xie, & McPherson, 2015; R Core Team, 2015).

PredAHT can assist professionals when evaluating possible cases of AHT and contribute to decision-making at multiple stages of the assessment pathway, by demonstrating how combinations of clinical features combine to estimate different probabilities of AHT. A crucial aspect of CPT development is determining its acceptability to the population for whom it is intended (Stiell & Wells, 1999; Reilly & Evans, 2006; Brehaut et al., 2010). Even valid and reliable CPTs may not be accepted or used (Stiell & Wells, 1999). If a CPT proves to be acceptable in addition to demonstrating a positive impact on patient care, its long-term and widespread dissemination and implementation would be justified; if not, the CPT could undergo modification and further evaluation (Brehaut et al., 2010). There is a need to better understand whether PredAHT is acceptable to child protection professionals and whether or not it is likely to be used in practice. Therefore this study aimed to explore the acceptability of PredAHT with a variety of professionals involved in the child protection process.

Methods

This was a qualitative semi-structured interview study. The study received ethical approval from the Cardiff University School of Medicine Research Ethics Committee (Ref: 15/35). This study is reported in accordance with the Consolidated Criteria for Reporting Qualitative Research (COREQ) guidelines (Tong, Sainsbury & Craig, 2007); a checklist is included in Appendix 1.

Participant Recruitment

Purposive sampling and snowball sampling were used to recruit participants for this study. We targeted clinicians, CPSWs, legal practitioners, police officers and pathologists involved in suspected AHT cases across south west United Kingdom (UK). A list of potential participants was identified through personal contacts of the research team and organizational websites. Personal contacts and organizations were sent an information sheet to explain the study and were asked to suggest suitable participants for interview. A random selection of individuals from each professional group were then invited to take part. We recruited participants with different levels of child protection experience and seniority (Figure 1). Individuals were contacted via email, with the exception of judges who were sent formal letters of invitation. In this study the term "clinician" refers to medical doctors and specialist nurses, who were sampled from three teaching hospitals and two district general hospitals across a range of specialties including pediatrics, radiology and neurosurgery. Most participating clinicians were experienced consultants or associate specialists, two were trainee doctors and one was a nurse. Senior CPSWs were team managers, who had a greater number of years of child protection experience than their junior counterparts. Senior police officers were inspectors or chief inspectors, and junior police officers were constables or sergeants. Judges had the greatest legal seniority with the responsibility of delivering the judgment on the evidence submitted by barristers or solicitors, and forensic pathologists had more experience of child protection investigations than the pediatric pathologist.

Interview Schedule Development

The semi-structured interview schedule explored three main themes 1) participants' evaluations of PredAHT 2) participants' opinions about the optimal way to present the calculated probabilities 3) participants' interpretation of probabilities in the context of

suspected AHT. The schedule was developed by LC and MF and revised following discussion with the research team (Appendix 2). Questions were derived following a review of the scientific literature on the identification of AHT, and the acceptability and evaluation of CPTs. The schedule was piloted with a police officer and a clinician, regarding the length, appropriateness, and content, and amended accordingly. The schedule included core open-ended questions, prompts and clarifying questions. It was a guide rather than a definitive list, to allow exploration of additional topic areas that might be raised by participants. Early responses influenced questions asked in later interviews; the schedule was updated as data collection and analysis progressed and new topic areas were raised.

Data Collection

Interviews were conducted by LC, a PhD student with training in qualitative research methods and qualitative interview techniques. No relationship was established between the interviewer and participants prior to the study. Informed consent, and permission for audio recording for verbatim transcription were obtained before each interview. Participants received a demonstration of the computerised PredAHT and a brief description of its development and purpose. Clinicians saw a version of PredAHT that allows them to input their own "pre-test" (prior) probability of AHT, based on factors other than the six included features that are pertinent to each case e.g. purported history, clinical presentation or psychosocial features. The prior probability influences the post-test probability of AHT provided by PredAHT, and is otherwise set at 34%, the prevalence of AHT in the data used to derive the tool.

When two participants declined to be audio recorded the interviewer made detailed notes of their responses. These were sent to the participants to check that they were a fair reflection of their views. The schedule was delivered to individuals, or at two small group

interviews (of three and five participants from the same professional group) where personal interaction between the participants was minimised, to maximise individual contributions from participants. Interviews lasted about 45 minutes and took place at the participants' workplace between June 2015 and September 2016. MF was present to record relevant field notes such as participant non-verbal behaviour and response to the interview, and critical reflections about the interview. No repeat interviews were conducted. In the interests of reflexivity, the interviewer considered how her own values and assumptions as a student involved in developing PredAHT might influence the interviews or the interpretation of the findings and a reflective journal was kept in an attempt to identify and minimize potential bias. It is acknowledged that the interviewer had a vested interest in PredAHT, hoping that it would be useful for the participants, and that they may have found it difficult to receive criticism or negative opinions about PredAHT. In early interviews with clinicians, the interviewer was acutely aware of her status as a psychology graduate, with no medical or legal training, but nevertheless conducting research in a medico-legal topic, and how this may affect the power relationship between the researcher and the participant. To break down power imbalances, every effort was made to build a rapport with the participants, and ensure that the interview was guided by them but stayed on-topic. The researcher made sure not to ask any leading questions, or impose their own views on the participants.

Data Analysis

Data analysis began shortly after the first interview using thematic analysis (Braun & Clarke, 2006). A general inductive approach enabled the results to be guided by the aims and objectives of the research, and the raw data (Bryman & Burgess, 1994; Dey, 1993). The Framework Method was used to manage, summarise, display, and synthesise the data and to facilitate analysis (Gale, Heath, Cameron, Rashid, & Redwood, 2013). Analysis followed

seven phases: transcription, familiarisation, coding, developing an analytic framework, applying the analytic framework, charting data into framework matrices, and interpretation (Gale et al., 2013). Initial codes were generated independently by LC, MF and HQS. These were jointly grouped into clearly defined categories that were further arranged under the three main overarching themes explored in the interviews. Discrepancies between coders were resolved by discussion and consensus. This process was undertaken in an attempt to minimize individual biases; 38% of the transcripts were independently double-coded. The joint analysis enabled the development of a preliminary analytic framework. Transcripts were imported into NVivo (QSR International Pty Ltd. Version 10, 2014), to organise and manage the data and assist with data analysis. Quotes pertaining to each category were retrieved and 'charted' into thematic framework matrices. Interviews ceased when thematic saturation was achieved within each group of participants (clinicians, CPSWs, police officers, legal practitioners and pathologists), which was verified using the constant comparative method (Glaser & Strauss, 1967). The final phase of the analysis involved abstraction and interpretation of the data. Participants were not asked to provide feedback on the study findings. Theme 1 is entitled 'participants' evaluations of PredAHT'. These were categorized as 'potential benefits of PredAHT', 'potential risks of PredAHT', 'provisos for the use of PredAHT ', 'use of PredAHT in court', and 'clinicians' views about the practical use of PredAHT'. Theme 2 is entitled 'participants' opinions about how to present the calculated probabilities'. These were categorized as 'percentage probabilities versus broad risk categories', 'confidence intervals' and 'additional suggestions'. Theme 3 is entitled 'participants interpretation of probabilities in the context of suspected AHT'. These were categorized as 'threshold criteria', and 'comments about PredAHT scores'. The systematic synthesis of the data excerpts into thematic matrices enabled a final inspection of the categories across cases, to identify

subcategories, i.e. the range of different elements being described under each category. All subcategories and their definitions are detailed in the analytic framework (Appendix 3).

Results

Participant demographics and response rates are shown in Table 2 and Figure 1. Fifteen of twenty-five (60%) clinicians (six had previously used a CPT), 1/4 (25%) judges and 2/4 (50%) pathologists were familiar with and overall had a positive opinion of CPTs; none of the CPSWs, barristers, solicitors, or police officers were aware of them. Data are presented using quotations, selected as examples of the themes and categories that emerged from the data. Within the quotations, square brackets represent text inserted for clarification. Word repetitions and irrelevant sections were removed and denoted by '…'.

Theme 1: Participants' evaluations of PredAHT

Potential benefits of PredAHT

PredAHT would be useful to support decision making as it is not influenced by personal feelings or opinions and could help reduce subjectivity in the assessment of risk.

"I think they're a good idea because they can be completely evidence based, so it takes all your feelings out of it because it's…child protection, there's lots of emotions." Clinician 9

"It would be helpful if a medical professional would have some confidence in saying it's an 85% chance because we would all understand what the chances were, because sometimes at strategy discussions you might get a pediatrician who will give an opinion, but as we all know, we take in messages in a different way. I might go back and record it in some way, the [child protection] social worker might go back and record it in a slightly different way." Police Officer 3

PredAHT would be helpful for heightening awareness of AHT, reinforcing or

increasing concerns about possible AHT.

"Where the number sits would help us to articulate that suspicion and perhaps work as a bit of a check. Probably in pushing it up, and highlighting to some people actually, you should be more suspicious because this is really unusual to get this combination." Police Officer 1

PredAHT could work both ways, encouraging participants to consider the possibility of non-abusive injury if PredAHT gave a low probability score for AHT.

"It would also be helpful for us not to panic too much in the sense of there is the idea of accidental head injury as well. So it's also helpful for us to take a step backwards and not think right it's abusive trauma." CPSW 1

Clinicians, CPSWs and police officers said that PredAHT would provide them with reassurance or confidence that their concerns, suspicions or investigations were justified and that it would be useful to support their professional opinions.

"I think as you used it more it would give you more confidence that actually, yes this is confirming that my level of suspicion is appropriate for the case…It would give you more confidence in making those decisions clinically." Clinician 15

The majority of clinicians would use PredAHT to back up their clinical opinion rather than to direct their decision-making.

"I wouldn't use it just to sway my opinion, but if I had an opinion of whether it's abusive or not, and then, using this validated tool, with the injuries found and the presence of head injury, it is likely, so that helps back up your opinion and hopefully then would add more weight to what you're saying." Clinician 21

However, some would find the score helpful if it did not agree with them.

"If there was a mismatch between my clinical opinion and the risk assessment tool that would cause me to stop and think and seriously consider whether I have gone down a bit of a blind alley with this and whether I need to stop and think again…It would be helpful just to

reassure us that we are doing the right thing or maybe to cause us to stop and think actually perhaps we shouldn't walk away from this one." Clinician 10

PredAHT would be useful for explaining, justifying or rationalizing decision-making in suspected AHT cases.

"Family courts, criminal courts might want to know how have you arrived at this decision and if I was asked, well, these are the facts....and I've documented clearly why I've made a decision about something. Any tool I think that helps gives some…statistical interpretation for police, for social services, for the medical professionals, it's robust and trustworthy I suppose then, I only see that as a good thing because we're all accountable for the decisions we make." Police Officer 7

Clinicians and pathologists suggested that PredAHT may help to standardize or modify the clinical assessment of suspected AHT cases by prompting clinicians to perform investigations such as a skeletal survey or ophthalmoscopy in line with international standards, and to review the results of investigations already undertaken.

"Is this patient a patient that may have been abused and if so [the tool] triggers safeguarding procedures for siblings and it flags up this is a child who is going to need an ophthalmologist to look in her eyes, a skeletal survey, and a child protection pediatrician on call. If that triggers all of those things to take place that would be great." Pathologist 1

"We would just do all those investigations on anyone under 1, but it's in that 1-3 [age group], where you're just that bit more unsure, whereas should we be doing these things and it might actually guide us." Clinician 1

Pathologists and legal practitioners, including judges, could also appreciate the value of PredAHT for advocating further investigations, even if they would not find it useful themselves.

"I think if it can be used to ensure that front line clinicians are actually encouraged to undertake exploration of what they've got at a better level then we'd be saying yippee absolutely that's the best that you can do for us…If you can get clinicians to actually do what they should be doing when they should be doing it and triggering the protocols that need...because we see quite a lot of missed opportunities with initial investigation and you can't go back and do it again." Judge 2

Similarly, police officers and CPSWs said that PredAHT may help to justify further action within their respective agencies.

"If we've got a figure that says actually there's an 80% chance that there's abusive trauma, then that child isn't safe at all…we need to be taking pro-active action and that would I think be supportive." CPSW 3

PredAHT could contribute to 'the bigger picture', as part of a wider information gathering process. Many described PredAHT as a useful addition to the 'toolbox', or 'a piece of the jigsaw puzzle'.

"It would form one part of your prosecution case wouldn't it? It wouldn't be an enormous part but it could form a part of the evidence you'd built generally…I can't see for a minute that it wouldn't be useful." Police Officer 4

PredAHT could be used at multi-agency meetings or as part of information sharing **to** facilitate communication about the likelihood of AHT.

"It would be valuable for talking to the police, [child protection] social workers... just to say, 'Listen we've got this…' Because they will always say to you, 'Is there anything else it could be? Are we getting this wrong? Are we missing something medical?' I think when you're able to say with a degree of certainty, 'No, this is what it is because this is a validated tool. With this combination of injuries this is how confident we can be' I think it is going to be valuable

for them as well...and you share it in the strategy meeting that would be very useful for me." Clinician 5

PredAHT would be useful for peer review or training. CPSWs in particular thought it was helpful to know that the six clinical features included in PredAHT are potential indicators of AHT.

"I think it will be very good for all child protection social workers dealing with these to know about these six things." CPSW 1

PredAHT may be most beneficial for so-called 'gray' cases, where there is considerable uncertainty surrounding the diagnosis, and most beneficial for those with the least experience in child protection.

"*You get those ones where you think 'this is really not likely to be this' but you've got to go through the steps, and the ones where there clearly is likely to be a problem, so it's those gray ones in the middle where this might come in more useful than the clear cut ones." Clinician 1*

"It might be helpful for someone who's never done pediatrics before and doesn't have the experience and the benefit of having done child protection work before and knowing these things…I think it is helpful for a very specific group of people." Clinician 20

Potential risks of PredAHT

Professionals may be over-reliant on PredAHT when making decisions in suspected AHT cases.

"If there was too much reliance placed on it at the beginning of an investigation and someone with little knowledge simply populated those fields present, absent, and came out with a low probability you know 14% or whatever and decided not to investigate, regardless

of the presence of other factors not in your fields, then that would be foolish and dangerous, so the tool itself is not a disadvantage, it's how its uses could be." Police Officer 6

"I think there's a potential for people to make it the single most significant part of the decision making process so we'd end up sat in meetings and people would ignore most of what I said and say 'What does the tool say? Oh 67%, right that's the decision made'. That would be my worry that people would over-use it or overstate its importance." Clinician 12

A low score could instil a false sense of security and appropriate investigations might not be carried out in the face of a low probability score.

"*If somebody is uncritically using this tool and they have got a child with an intracranial head injury, and head and neck bruising they are not worried then because the score says that it's less than 15% then that would be an incorrect use of this tool." Pathologist 1*

However, reassuringly, all participants said that they would still carry out appropriate investigations if they received a low score from PredAHT, as there may be other features of the case that warrant further enquiry.

"If it was low like that, but there were previous allegations of abuse, dad had a violent background, that sort of stuff. That would sway me towards being quite concerned about this. So, I guess it's about the attendant circumstances around it. So if it was low, it would help me, but I would still look at the bubble around it and what's going on." Police Officer 3

Concern was expressed that PredAHT wouldn't be used as intended, alongside other known information about each case, and they agreed that it should never be used in isolation from other factors. Several clinicians and pathologists said that PredAHT was too reductionist and crude, comparing it to a box-ticking exercise.

"My main concern is people not taking into account the history or the other facts because they've got a big number on this score." Pathologist 2

The accuracy of PredAHT was questioned. The sensitivity and specificity that participants would be willing to accept was discussed, together with the implications of false positive and false negative predictions.

"The key thing is in how many cases is this wrong? And if it's wrong in any, then you've just got to ask yourself is this safe?" Pathologist 4

"Hopefully this will help us find all cases of abusive head trauma but there is a chance that we might label a non-abusive one as an abusive one as well, but I would probably weigh the benefit more than the risks, because if this is helping me to identify the really vulnerable children, I would rather use it…as long as I am protecting the vulnerable children, I would find it useful." Clinician 22

A few pathologists and clinicians, including a neurosurgeon, stated that they do not need a tool to make decisions about suspected AHT, and would therefore not use PredAHT.

"That's what we do in our brains, we put all the information together and spit out the probability based on our experience." Clinician 19

"It's something that is an irrelevance to me in that, one might take the view that this is an attempt to make my task less onerous by placing in my mind the conclusions of others, or their interpretation of the evidence, when it is my role to look at that evidence and the literature myself. This I fear might be regarded as a substitute for individual thought." Pathologist 4

Two clinicians were unsure about how much PredAHT would add to the investigative process, and could not say whether they would use it in practice or not.

"I have to say my initial thought looking at it is I'm not sure how much more it would add if you've done all the investigations already." Clinician 13

Some participants thought that there are important clinical and historical features missing from PredAHT, e.g. skull fractures, bruising patterns, spinal injury, or a history of trauma, and questioned why they were not included.

"You don't have a history of extraordinary trauma as an option. So the other thing is a non-declaration of the history that would be massive wouldn't it. Or no history of any injury…just woke up and the baby was like that. That would be a massive predictor I would imagine." Clinician 24

PredAHT cannot take into account specific details of the clinical features, including the severity of injury, injuries of different ages, and the precise locations and patterns of injuries, some of which may be highly specific for AHT.

"*What about a healing fracture as opposed to a recent fracture and position of the fracture, and particularly rib fractures are they at the front of the chest of a child who has had resuscitation or are they posterior ribs, and it's this granularity that we are grilled on and we have to take into account when we are giving our overall opinion but for a quick and dirty or 'Should I be contacting child protection services, should I be contacting our pediatrician on call for child protection?' This sort of thing is great." Pathologist 1*

"Not all RHs are the same, you can have one in one eye and five in the other but if they're not in the layers that you'd expect them to be…you'd have to make some allowance for, not only the categories but subcategories of that...It needs to be more refined." Judge 2

Some pathologists and judges said that PredAHT may condition their decision-making or inadvertently introduce bias into the decision-making process.

"You would almost make it more difficult for the judge because the judge would then have to disentangle the expert opinion from either an apparent bias or an unconscious bias that might be established by the fact that the expert had looked at the clinical tool." Judge 3

Provisos for the use of PredAHT

Many participants would only use PredAHT with a proviso; e.g. alongside their professional judgment, with more information about the definition of the six features, if it was kept up to date, with knowledge of the quality of the data on which PredAHT is based, with an understanding of how it works and how it is to be used, after agreeing acceptable risk thresholds with multi-agency colleagues, and if it was accepted by their colleagues (Table 3).

Use of PredAHT in court

Most clinicians, CPSWs and police officers thought that PredAHT would be useful in court, because it is evidence-based and validated. CPSWs in particular felt that it would be useful in the family courts for future safeguarding of children, where the standard of proof is based upon the balance of probabilities.

"*In the court arena I think it's going to be really very useful because it's not our hunch against the next doctor's hunch, you know? And I think people's general opinion that babies aren't injured by their carers and their parents...people don't want to hear that and they certainly don't want to believe it and acknowledge that this is happening, but if you've got a validated tool saying, 'Actually this is what has happened to this baby because of the other injuries that we've seen' then I think it's going to be very valuable indeed." Clinician 5*

"It helps when going to court with the balance of probabilities if you can prove over 51%, that's the number I have in my head…that's what we've got to convince evidence of a judge of." CPSW 2

However, pathologists and legal practitioners, including judges, expressed caution regarding its use in court, particularly in the criminal courts.

"What the criminal standard which is beyond reasonable doubt would make of that, because the decision in these circumstances would be that of a jury, again huge caution in thinking

this is effectively steering a jury into saying it's 85%, it's beyond reasonable doubt therefore we've got no choice but to convict." Legal Practitioner 1

Some felt that PredAHT would be irrelevant because it cannot account for every detail of every case, and each case must be considered based on the entirety of the evidence.

"As lawyers we would probably want to treat it with extreme caution. Simply because this tool cannot cover every factor in every situation that we have to deal with." Legal Practitioner 1

Others remarked that PredAHT may not stand up under cross-examination, or that the defence will claim that their case falls into the reverse probability of non-AHT given.

"You would have to prove the tool in every case. You'll be cross examined about how it's been put together, how you've weighted the factors. There's always something that somebody can find if you're really trying to pull something apart. Then it goes out the window really evidentially." Judge 2

"We would be arguing well why isn't this one in the 15% of cases that suggests that it isn't non-accidental?" Legal Practitioner 1

In addition, some participants pointed out that PredAHT will not help to identify the perpetrator in suspected AHT cases.

"That doesn't help us with who caused it, it just says 'what's the probability of it being an abusive trauma' so there is that other element we have to consider." Legal Practitioner 1

Some clinicians and judges discussed historical child protection court cases that involved the use of statistical evidence, and the impact and implications of such cases on the subsequent acceptance of statistics in the courtroom.

"A slight worry any pediatrician will have, a study putting statistics up like that, is the way that [Roy] Meadow [UK pediatrician] was chopped down with statistics." Clinician 17

However, one judge thought that PredAHT would definitely be useful, in both the family and criminal courts.

"It will help to remind the courts and the experts that a certain combination of features does make abusive head trauma a more likely explanation...I think it would have the same role in the criminal courts. Even though the standard of proof is different, it would still be useful at the fact finding stage." Judge 4

Despite their reservations, the majority of the legal practitioners interviewed, including judges, agreed that they would probably take the PredAHT score into account if it was included as part of a medical court report.

"The way the courts see these matters from a child protection point of view is an analytical approach where you need the best evidence possible. If this is something that feeds into a medical report, by an expert who understands it, then I'm delighted to have that. Hopefully it can help to make the right decision for the family, because it is life changing." Judge 4

Clinicians' views about the practical use of PredAHT

All thought that PredAHT is simple to use, and not too time-consuming to complete.

"Time is always a disadvantage in getting people to fill these out sometimes, but this is relatively simple and straightforward so I don't imagine it being a huge issue." Clinician 14

The majority thought that PredAHT would be most useful for inpatients admitted to a ward or Pediatric Intensive Care Unit (PICU), and less useful in the Emergency Department (ED), where information about fractures and RHs is unlikely to be available. However, one emergency medicine pediatrician thought that PredAHT may have a role in the ED to prompt an initial referral to the safeguarding team.

"If they were unwell enough to go to PICU you may use that tool much less in the ED. If they were somebody that was going to a ward then you would probably use it more. I think it would depend on the patient and how sick they were." Clinician 16

There were different views regarding the stage of the assessment process that PredAHT would be most useful. Some would only use it once all relevant investigations were completed, to assist with report writing or reaching their final conclusions.

"In my opinion, there's not much point in using it if you have too many unknowns there...I personally would certainly like it for when I have to write my report." Clinician 7

However, others could see the value of PredAHT at multiple stages of the assessment process, and would use it more than once during a case to support their decision-making.

"I would probably use it as soon as I knew about the case, just to give me some idea, and then as more data is collected you could add it in and see how it changes your figure, and then you've got your last kind of figure then is what is going to be the most important one right at the end." Clinician 1

PredAHT could be completed by general and community pediatricians and intensivists, but most agreed that it should be completed by a consultant. One clinician thought that it should be a team exercise.

"Whether it would be something that would be used by the lead consultants in PICU or a general pediatrics consultant where they are thinking do we need to get the safeguarding team involved or not and then potentially I suppose used by a safeguarding consultant when it came to writing up." Clinician 8

Clinicians reflected on how PredAHT could be integrated into the clinical workflow and implemented in clinical practice. Although they acknowledged that each hospital has its own way of working, most thought that it would not be too difficult to incorporate the computerised version of PredAHT into existing hospital intranet systems. Some suggested

including reminders or specific references to PredAHT on existing departmental or safeguarding paperwork.

"It could well go on to the intranet as an app…and maybe a reference to it as a little reminder on our safeguarding proforma." Clinician 4

Some clinicians could see the value of including their own prior probability score in the calculation, but the majority felt that this would introduce too much subjectivity into PredAHT and that they would need guidance on how to use this element of the tool.

"I think that's an important element to bring in because a lot of our decisions are often based around the history and does the history fit, is it consistent, stuff about whether they're presenting late, stuff about what the family background and social history is." Clinician 13

"I'd feel happier with the six features on its own, because I know that's very evidence based, isn't it, so that's fine. I do have an issue with the prior probability without some objectivity around it, because it's easy to think oh well, a child on the Child Protection Register and that ups my concerns and whereas a middle class family doesn't, and I think it is something you need to be really objective about." Clinician 3

Theme 2: Participants' opinions about the optimal way to present the calculated probabilities

Precise percentage probabilities of abuse were preferred, rather than broad risk categories such as low, medium or high likelihood of abuse. While some suggested presenting both, they were unsure as to what percentages would equate to low, medium and high. Some participants felt that confidence intervals would be unnecessary, however others felt that they would be an important addition. Additional suggestions included background information/data about PredAHT, disclaimers, and visual aids (Table 4).

Theme 3: Participants' interpretations of probabilities in the context of suspected AHT

Threshold criteria

Although all participants maintained that they have a very low threshold for suspicion of AHT in young children with intracranial injuries, their probability thresholds for suspecting abuse varied widely. Some participants would only feel confident to completely rule out AHT if the percentage probability was less than 1%, while others had higher thresholds.

"If it's something like in the thirties, gosh that's a really hard thing to factor in isn't it then, in terms of decision making, it still sounds quite high to me. You almost want it to be a 0.1% chance of it being an abusive head trauma to feel confident in your decision, because even at 30%, that's like one out of three families, that was abusive isn't it?" CPSW 9

"If that said to me there's a 1% chance then there's still a 1% chance. It's helping me it's not telling me there's no chance is it?" Police Officer 8

"I want to say sometimes I'm not happy about not taking any further action at times, but I'd have to be I don't know maybe 20%?" CPSW 2

Many participants simply could not put a figure on their threshold for abuse, stating that if there was any chance at all that it could be, then they would investigate further, and commenting that each case is dependent on the attendant circumstances around it. One CPSW indicated that often her risk judgments are very different to her colleagues', highlighting inconsistencies within the assessment process.

"I don't think I can put a figure on my threshold because it depends sometimes I look at something and think 'Why are we going out on this?' And then something else, 'We should've looked at that, why did we have all of this and we haven't done anything with it?'" CPS W 4

Participants also had different perceptions of what the expressions 'low', 'medium' or 'high' likelihood of abuse might mean in percentage terms.

"Less than 50 would be low. Maybe 50–70 medium and then over 70 high." CPSW 6

"I suppose low, medium and high can mean anything can't it, I'm guessing its maybe up to 30%, 60%, 90%." CPSW 8

Comments about PredAHT scores

After receiving a demonstration of PredAHT, participants offered their opinions on the probability scores that it gives for different combinations of features. They were told that intracranial injury with head/neck bruising alone gives a score of 14.7%. Some participants thought this score was low. Others interpreted the score to be an unacceptably high level of risk, while others still thought this figure could be interpreted in both ways.

"I think that's really low 14.7%." CPSW 4

"14.7%, what do I think of that? I still think it's bloody high." Police Officer 1

Participants were uncomfortable that PredAHT can give scores at or very close to 100% for certain combinations of features.

"99.6% and I've ticked rib fracture present, head and neck bruising present, apnoea present, seizure present. That I find hard you're saying essentially that's definitely abusive head trauma…I agree, I'd be very worried if I had that combination of features but I wouldn't say it's nearly a hundred." Pathologist 2

Others stated that PredAHT scores too low for certain combinations of features.

"So if you're fitting and have subdurals but you don't have RHs or any other markers, it's saying it's not non-accidental injury. I would be a lot more hawkish than that." Clinician 24

"I think that my findings were stronger than that score…the RH was very strong, in all layers…I might've hesitated if I'd seen 57%." Judge 2

Discussion

The study findings suggest that PredAHT would support the decision-making of clinicians, CPSWs, and police officers investigating suspected AHT, and provide them with greater confidence in expressing their opinion in the child protection and court setting. Benefits were perceived by junior and senior practitioners with different levels of child protection experience, and across all specialities with the exception of a neurosurgeon, although it was acknowledged that PredAHT may be most useful for those with the least child protection experience. Pathologists and legal practitioners, including judges, thought PredAHT to be useful as a screening tool for ruling in further clinical or multidisciplinary investigations, however with the exception of one judge, they expressed caution regarding its use in court. In practical terms, clinicians found PredAHT to be simple to complete and thought it would be straightforward to implement into existing hospital systems.

All professionals who come into contact with children and families have a duty to safeguard children and young people and should receive regular training to ensure that they are competent in their respective roles (Department for Education, 2015). Two UK government reports on social work interventions required in the child protection arena stated the need for the development of an evidence-based approach and learning culture, to inform good practice; one explicitly recommended the use of standardized tools to support decision-making and analysis of information about whether a child is suffering, or likely to suffer significant harm (Barlow et al., 2012). The other highlights the importance of improving the skills and knowledge of CPSWs (Department of Education, 2014). Our study highlighted gaps in the training and knowledge of professionals working in child protection; many were unaware that some of the clinical features included in PredAHT were indicators of AHT. This is consistent with a recent study that demonstrated gaps in knowledge and training about bruising amongst CPSWs (Matthews, Kemp, & Maguire, 2017). Improved communication within and between agencies is critical for identifying patterns and preventing further injury

(Joughin, 2003) and has been recommended in the UK Safeguarding Children Research Initiative report (Davies & Ward, 2012). The current study confirmed that PredAHT would facilitate interagency communication about the likelihood of AHT.

Clinicians stated that PredAHT would give them more confidence in expressing their opinions about the likelihood of AHT in their court reports and in court settings. While pathologists and legal practitioners, including judges, appreciated the value of PredAHT for encouraging standardization of clinical investigations, and further clinical or social assessment, overall they expressed caution regarding its potential use in court. However, the majority would incorporate the probability score with all of the other evidence if it was provided as part of a medical report and PredAHT was accepted by the medical community. Previously, high profile cases involving the misinterpretation of statistics and probabilistic evidence in the courtroom have caused controversy (R *v* Clark, 2000; R *v* Cannings, 2004) and led to the development of a working group within the Royal Statistical Society in the UK, to improve the use of statistics in the administration of justice. They recommended a broad programme of education for judges, lawyers, and expert witnesses in probability theory and statistics (Aitken, Roberts, & Jackson, 2010). One judge and one pathologist were concerned that PredAHT would introduce cognitive bias into their decision-making, suggesting that they may not trust PredAHT to be a valid piece of evidence in their decision making.

The scientific literature confirms that there are specific patterns of intracranial injury, haemorrhagic retinopathy (Wright, 2017), and spinal injuries associated with AHT (Choudhary, Ishak, Zacharia, & Dias, 2014). Various psychosocial variables may also be influential (Pierce et al., 2017). Some participants wanted these additional features incorporated into PredAHT. However, a massive multi-center prospective study would be needed in order to add further variables. Meanwhile, the six clinical features in PredAHT are easily identifiable in the early phase of clinical assessment.

Participants had varying opinions about what percentage probabilities equate to the terms ‘low’, ‘medium’ or ‘high’ likelihood of abuse, and diverse probability thresholds for suspecting abuse, although participants stated that PredAHT helped them to quantify risk. This is consistent with previous studies that demonstrated that pediatricians struggle to define “reasonable suspicion” or “reasonable medical certainty” of abuse (Dias, Boehmer, Johnston-Walsh, & Levi, 2015; Levi & Brown, 2005). Thresholds in child protection social work have been the subject of much debate in recent years and are affected by a wide range of organisational factors, relationships with other professionals, and individual biases, heuristics and value systems (Platt & Turney, 2014). Similarly, participants postulated that PredAHT may be most useful for ‘gray’ cases, where there is significant uncertainty surrounding a diagnosis of AHT. In reality, the interpretation of the PredAHT score will depend upon individual perception of risk. PredAHT is designed to be an assistive tool rather than a decision rule, which typically recommends a direct course of action based on the results (Reilly & Evans, 2006); PredAHT provides no recommendations for professionals on what to do based on specific scores. Despite this, the majority of participants thought PredAHT would be useful for supporting their opinions and decision-making.

Previous research has identified barriers to the use of CPTs, some of which emerged from the interviews, such as scepticism of “cook-book” medicine, belief that clinical judgment is superior to the tool, distrust of the accuracy of the predictors and concern that the CPT does not address all relevant factors (Reilly & Evans, 2006). Reilly and Evans (2006) offer a number of strategies to overcome these barriers, including comparing clinical judgment with the CPT, and checking whether any excluded factors affect the CPTs predictions. Skull fracture was analysed within the original derivation study and did not

discriminate between AHT and non-AHT (S. Maguire et al., 2011). A study comparing PredAHT with clinical judgment is underway.

The findings have implications for the further development and implementation of PredAHT. Given that some participants wanted confidence intervals whilst others did not, it may be sensible to include an option to display these. Participants' identified a number of conditions under which they would use PredAHT, namely if it was accepted by their colleagues, used alongside their professional judgment, and if they understood how it works. These are consistent with a study exploring the acceptability of a tool to identify abusive or neglectful burns (Johnson, Hollén, Kemp, & Maguire, 2016). Any training on PredAHT would need to encompass these elements. Over-reliance on PredAHT, concern that it may be used improperly or failure to investigate appropriately if a low score is given, and the potential ramifications of "false positives" or "false negatives" emphasise the importance of providing clear guidance to practitioners about how PredAHT is intended to be used, namely as an assistive CPT, and *not* a diagnostic tool. Finally, PredAHT allows clinicians to factor in aspects of the history and other risk factors by incorporating their own prior probability of AHT into the calculation. Clinicians felt this element of PredAHT was subjective, and were unsure whether they would be comfortable estimating a prior probability of AHT in light of potential racial and socioeconomic bias (Wood et al., 2010). This suggests a lack of knowledge amongst clinicians of the evidence base regarding psychosocial risk factors for AHT. A recent survey found that less than half of health care professionals are adequately trained or prepared to identify risk factors associated with maltreatment (Foster, Olson-Dorff, Reiland, & Budzak-Garza, 2017). Training on how to estimate a prior probability of AHT should be incorporated into full implementation, as requested by clinicians.

Strengths and Limitations

There are many derived CPTs for children; few are validated, while virtually none undergo impact analysis (J. Maguire et al., 2011), and very rarely do investigators determine the acceptability of CPTs prior to their use. To our knowledge this study is one of only two studies (Johnson et al., 2016) exploring the acceptability of a CPT developed for use in child protection, and the first study to have done so with a wide range of professionals. Strengths of the current study include the rigorous data analysis methods employed, and the depth, detail, and richness of the data collected. Semi-structured interviews enabled the interviewer to build rapport with the participants (Fontana & Frey, 1994), and led to richer data than might be gained from a more structured approach, a survey or questionnaire. PredAHT is the only CPT that we are aware of that can estimate a predictive probability of AHT given different combinations of multiple clinical features, at various points along the assessment pathway. Other CPTs have been developed for use in the ED to assist clinicians deciding which high-risk children should undergo computed tomography (Berger et al., 2016), and for the PICU to help exclude AHT when negative (Hymel et al., 2014). However, whether clinicians or other practitioners would be prepared to use these CPTs in practice is unknown.

Despite the inclusion of a range of professional groups and clinical subspecialties, additional groups could have made valuable contributions, for example neurologists, intensivists, or ophthalmologists. Most of the clinicians were consultants working in teaching hospitals, and less experienced clinicians, and radiologists, neurosurgeons, and nurses, were under-represented. However it should be noted that probabilistic representativeness is not a goal of qualitative research (Popay, Rogers & Williams, 1998). The fourth pathologist interview revealed some new insights that the other pathologists had not brought up, suggesting that data saturation may not have been reached within this professional group. This does not necessarily invalidate the findings for this group but rather means that further exploration of the topic may be warranted with these professionals (O'Reilly & Parker,

2012). Given the sensitive nature of the research, some respondents may have shown a social desirability bias and responded in a manner likely to be viewed as favourable by the researcher, however a number of participants exhibited an unfavourable view of PredAHT and were open about their opinions and intentions not to use it. Finally, qualitative research inevitably relies on the researcher's interpretations, however, subjective bias was minimized by using three trained qualitative researchers to double-code the data and resolve disagreements through discussion and consensus.

Conclusions

This evaluation has demonstrated that PredAHT is acceptable to child protection professionals across a range of disciplines assessing suspected AHT cases, and that they would be willing to use it as an adjunct to their decision-making. Although it may be most useful for those with the least child protection experience or knowledge, it is applicable to all professionals working in this area as it may help to reduce missed cases of AHT. These results confirm that the addition of a precise and objective evidence-based probability score that calculates the risk of AHT for child protection professionals is acceptable and potentially useful. This tool, when used in conjunction with a full clinical and social history, has the potential to standardize clinical assessment, and minimize subjectivity when weighing up the clinical features in cases of possible AHT. Feasibility work is underway to determine whether it is possible to evaluate the impact of PredAHT when it is applied in clinical practice. This will inform the planning and design of a formal impact analysis study and a long term implementation and dissemination plan to maximise uptake (Stiell & Wells, 1999).

References

Aitken, C.G.G., Roberts, P., & Jackson, G. (2010). *Fundamentals of probability and statistical evidence in criminal proceedings: guidance for judges, lawyers, forensic scientists and expert witnesses*. London, UK: Royal Statistical Society.

Barlow, J., Fisher, J.D., & Jones, D. (2012). *Systematic review of models of analysing significant harm*. Oxford: Department for Education.

Berger, R.P., Fromkin, J., Herman, B., Pierce, M.C., Saladino, R.A., Flom, L.,…Kochanek, P.M. (2016). Validation of the Pittsburgh infant brain injury score for abusive head trauma. *Pediatrics, 138*(1), e20153756.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Brehaut, J.C., Graham, I.D., Wood, T.J., Taljaard, M., Eagles, D., Lott, A.,...Stiell, I.G. (2010). Measuring acceptability of clinical decision rules: Validation of the Ottawa acceptability of decision rules instrument (OADRI) in four countries. *Medical Decision Making, 30*(3): 398–408.

Bryman, A., & Burgess, R. (1994). *Analyzing qualitative data*. London: Routledge.

Canadian Paediatric Society. (2007). *Multidisciplinary guidelines on the identification, investigation and management of suspected abusive head trauma*. Ottawa: Canadian Paediatric Society.

Chang, W., Cheng, J., Allaire, J.J., Xie, Y., & McPherson, J. (2015). Shiny: Web Application Framework for R. R package version 0.11.1. http://CRAN.R-project.org/package=shiny

Choudhary, A.K., Ishak, R., Zacharia, T.T., & Dias, M.S. (2014). Imaging of spinal injury in abusive head trauma: a retrospective study. *Pediatric Radiology, 44*(9), 1130–40.

Christian, C.W., & Committee on Child Abuse and Neglect, American Academy of Pediatrics. (2015). The evaluation of suspected child physical abuse. *Pediatrics, 135*(5), e1337–e1354.

Cowley, L.E., Morris, C.B., Maguire, S.A., Farewell, D.M., & Kemp, A.M. Validation of a prediction tool for abusive head trauma. *Pediatrics*, *136*(2), 290–298.

Davies, C., & Ward, H. (2012). *Safeguarding children across services: messages from research*. London and Philadelphia: Jessica Kingsley.

Department for Education (2014). *Child protection, social work reform and intervention: research priorities and questions*. London: Department for Education.

Department for Education (2015). *Working together to safeguard children: statutory guidance on inter-agency working to safeguard and promote the welfare of children*. London: Department for Education.

Dey, I. (1993). *Qualitative data analysis*. London: Routledge.

Dias, M.S., Boehmer, S., Johnston-Walsh, L., & Levi, B.H. Defining 'reasonable medical certainty' in court: What does it mean to medical experts in child abuse cases? (2015). *Child Abuse & Neglect, 50,* 218–227.

Fontana, A., & Frey, J.H. (1994). Interviewing: the art of science. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 361–376). Thousand Oaks: Sage Publications.

Foster, R.H., Olson-Dorff, D., Reiland, H.M., & Budzak-Garza, A. (2017). Commitment, confidence, and concerns: Assessing health care professionals' child maltreatment reporting attitudes. *Child Abuse & Neglect*, *67*, 54–63.

Gale, N.K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013) Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology, 13*(117), 1–8.

Glaser, B.G., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Chicago, IL: Aldine Publishing Co.

Hymel, K.P., Armijo-Garcia, V., Foster, R., Frazier, T.N., Stoiko, M., Christie, L.M.,…Pediatric Brain Injury Research Network (PediBIRN) Investigators. (2014). Validation of a clinical prediction rule for pediatric abusive head trauma. *Pediatrics, 134*(6):e1537–1544.

Jenny, C., Hymel, K.P., Ritzen, A., Reinert, S.E., & Hay, T.C. (1999). Analysis of missed cases of abusive head trauma. *Journal of the American Medical Association, 281*(7), 621–626.

Johnson, E.L., Hollén, L.I., Kemp, A.M., & Maguire, S. (2016). Exploring the acceptability of a clinical decision rule to identify paediatric burns due to child abuse or neglect. *Emergency Medicine Journal, 33*, 465-470.

Joughin, V. (2003). Working together for child protection in A&E. *Emergency Nurse*, *11*(7):30–38.

Letson, M.M., Cooper, J.N., Deans, K.J., Scribano, P.V., Makoroff, K.L., Feldman, K.W., & Berger, R.P. (2016). Prior opportunities to identify abuse in children with abusive head trauma. *Child Abuse & Neglect*, *60*, 36–45.

Levi, B.H., & Brown, G. Reasonable suspicion: a study of Pennsylvania pediatricians regarding child abuse. (2005). *Pediatrics, 116*(1), e5–e12.

Maguire, J.L., Kulik, D.M., Laupacis, A., Kuppermann, N., Uleryk, E.M., & Parkin, P.C. (2011). Clinical prediction rules for children: a systematic review. *Pediatrics, 128*(3), e666–e677.

Maguire, S.A., Kemp, A.M., Lumb, R.C., & Farewell, D.M. (2011). Estimating the probability of abusive head trauma: a pooled analysis. *Pediatrics*, *128*(3), e550–564.

Maguire, S., Pickerd, N., Farewell, D., Mann, M., Tempest, V., & Kemp, A.M. (2009). Which clinical features distinguish inflicted from non-inflicted brain injury? A systematic review. *Archives of Disease in Childhood, 94*(11), 860–867.

Matthews, L., Kemp, A., & Maguire, S. (2017). Bruising in children: exploring the attitudes, knowledge and training of child protection social workers and the interface with paediatricians regarding childhood bruising. *Child Abuse Review*, *26*(6), 425-438.

O'Reilly, M., & Parker, N. (2012). 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research, 13*(2): 190-197.

Pierce, M.C., Kaczor, K., Acker, D., Webb, T., Brenzel, A., Lorenz, D.J.,…Thompson, R. (2017). History, injury, and psychosocial risk factor commonalities among cases of fatal and near-fatal physical child abuse. *Child Abuse & Neglect, 69*, 263–277.

Platt, D., & Turney, D. (2014). Making threshold decisions in child protection: a conceptual analysis. *The British Journal of Social Work, 44*(6), 1472–1490.

Popay, J., Rogers, A., & Williams, G. (1998). Rationale and standards for the systematic review of qualitative literature in health services research. *Qualitative Health Research, 8*(3): 341–351.

QSR International Pty Ltd. (2014). NVivo qualitative data analysis Software; Version 10.

R Core Team. (2015). R: A language and environment for statistical computing (Version 3.2.3). R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

R *v* Clark, EWCA Crim 54, Oct. 2nd, 2000.

R *v* Cannings, EWCA Crim 01, Jan. 19th, 2004.

Reilly, B.M., & Evans, A.T. (2006). Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Annals of Internal Medicine, 144*(3), 201–209.

Stiell, I.G., & Wells, G.A. (1999). Methodologic standards for the development of clinical decision rules in emergency medicine. *Annals of Emergency Medicine, 33*(4), 437–447.

Tong A., Sainsbury, P., & Craig, J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6):349–57.

van Buuren, S, & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67.

Wood, J.N., Hall, M., Schilling, S., Keren, R., Mitra, N., & Rubin, D.M. (2010). Disparities in the evaluation and diagnosis of abuse among infants with traumatic brain injury. *Pediatrics, 126*(3), 408–414.

Wright, J.N. (2017). CNS Injuries in abusive head trauma. *AJR American Journal of Roentgenology*, *208*(5), 991–1001.

Table 1. The six features included in the Predicting Abusive Head Trauma clinical prediction tool

Feature	Description
Head or neck bruising	Any documented bruising to head or neck
Seizures	Any documented seizures from a single seizure to status epilepticus
Apnea	Any apnea documented in the initial history or during inpatient stay
Rib fracture	Any rib fracture documented after appropriate radiologic imaging
Long-bone fracture	Any long-bone fracture documented after appropriate radiologic imaging
Retinal hemorrhage	Any retinal hemorrhage documented after indirect ophthalmologic examination by a pediatric ophthalmologist

Published previously in Pediatrics (Cowley et al., 2015) and reproduced here with permission

Table 2
Demographics of child protection professionals participating in an evaluation of the acceptability of the Predicting Abusive Head Trauma clinical prediction tool

	Clinicians (N=25)		CPSWs (N=10)		Legal Practitioners (N=9)		Police Officers (N=8)		Pathologists (N=4)	
	n	%	n	%	n	%	n	%	n	%
Gender										
Female	16	64	7	70	7	78	3	37.5	0	0
Male	9	36	3	30	2	22	5	62.5	4	100
Age group										
25–34	2	8	2	20	2	22	0	0	1	25
35–44	11	44	5	50	1	11	5	62.5	1	25
45–54	8	32	1	10	4	45	3	37.5	1	25
55–64	4	16	2	20	2	22	0	0	1	25
Ethnicity										
White British	19	76	10	100	8	89	8	100	4	100
White Other	4	16	0	0	1	11	0	0	0	0
Indian	2	8	0	0	0	0	0	0	0	0
Years in CP										
<5	0	0	2	20	1	11	3	37.5	0	0
5–9	6	24	3	30	1	11	2	25	2	50
10–20	7	28	4	40	4	45	3	37.5	0	0
>20	12	48	1	10	3	33	0	0	2	50
CP training										
Yes	25	100	10	100	3	33	7	87.5	4	100
No	0	0	0	0	6	66	1	12.5	0	0
Pediatric HI training										
Yes	18	72	1	10	3	33	4	50	3	75
No	7	28	9	90	6	66	4	50	1	25

CPSWs = child protection social workers, CP = child protection, HI = head injuries.

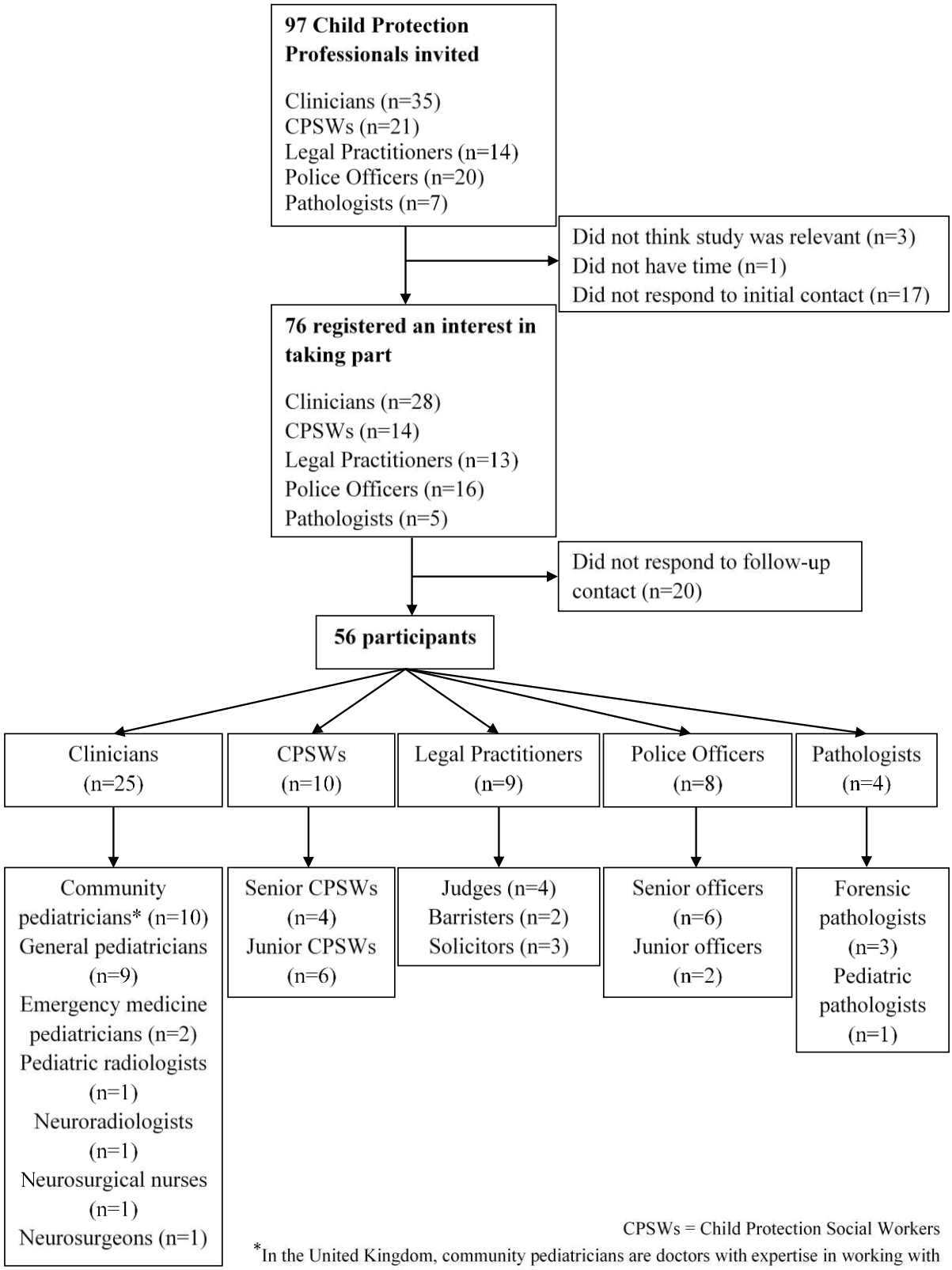
Table 3
Child protection professionals' provisos for the use of the Predicting Abusive Head Trauma clinical prediction tool

Alongside professional judgment	*"I guess it's probably a combination of that along with a bit of professional judgement tied in… if you look at it in combination with other professional opinion, what else you know, what information you found out, then it could inform part of that pool of information."* ***Police Officer 3*** *"Yeah I think it's useful, it should not stop you from thinking I think you should still think outside the box and not 100% rely on it but I think as an additional tool to your clinical decision making, I think it is supportive and helpful."* ***Clinician 7*** *"This in addition to the rest of our assessment is really, really helpful."* ***CPSW 1***
Definition of the six features	*"I would need a bit more information…you know I don't really know what retinal hemorrhages are."* ***CPSW 6*** *"Apnea, presumably you'd have a definition of how long that's for and stuff like that?"* ***Pathologist 2***
If kept up to date	*"There's issues of keeping it up to date, you can't just do it once and then not revisit it, can you?"* ***Clinician 3*** *"To keep its credibility it would have to evolve with current thinking, so it's a continual process isn't it?"* ***Police Officer 2***
Quality of the data on which PredAHT is based	*"I'd want to look at the original research and how the original cohort of patients were diagnosed with abusive head trauma and what's the robustness of that diagnosis in the first place, that the tool is then based on."* ***Clinician 12*** *"I would never use something like this without reviewing the publication and looking at the statistics and checking out that I was personally happy with the statistical analysis, because otherwise I'm just putting stuff into boxes."* ***Pathologist 2***
Understanding how it works and how it is to be used	*"I would have to understand it and be able to explain it in court, so I'd need to come and have a little training session."* ***Clinician 3*** *"You would need to explain the unknown parts of it…as well you could do a small tutorial based on four or five cases if people want to get experience on how to use it."* ***Clinician 14*** *"It's important to understand what informs the figure, because otherwise it becomes a checklist…I think people need to have an understanding of what the tool is and how it is to be used."* ***CPSW 5***
Agreeing accepted risk thresholds with colleagues	*"We can all have that figure and we can all explore then what that figure means to each independent agency, and what it means for that child and actually what safeguards need to be in place because of it."* ***CPSW 3*** *"I think within a team, there needs to be consistency as to what it's meaning at that point in time."* ***Clinician 2*** *"You want that consistency and agreement as to what the results could mean."* ***Police Officer 1***
If accepted by colleagues	*"It would be only useful for us if it's accepted by the medical profession."* ***Legal Practitioner 2*** *"Whether it be the safeguarding board or the child death overview panel, you'd want something where the social workers and the pediatricians and the police all come together and agree that this is useful."* ***Police Officer 1*** *"That would give me the most confidence really if the medical professionals were on board with it."* ***CPSW 6***

Table 4
Child protection professionals' preferences for the presentation of the calculated probabilities from the Predicting Abusive Head Trauma clinical prediction tool

Percentage probabilities versus broad risk categories	*"For a decision-making tool a percentage is spot on from my point of view, because I think everyone will understand it."* ***Police Officer 2*** *"A percentage just makes it a little bit more tangible, doesn't it, it's very real, because I find it difficult, I always have done, to quantify risk, so if it can be done for me, then yeah…"* ***Clinician 6*** *"I am happy with the percentage, I am confident with that because that is where we are at with safeguarding children, we are talking about the balance of probabilities and so we are talking anything over 50 per cent probability should lead to further action and further evaluation where you may still get information that shifts it the other way."* ***Clinician 10*** *"I'd personally prefer it if captured with likeliness and you might have very suspicious, strongly indicative or not likely."* ***Clinician 12*** *"I prefer it like that because low, medium, high can mean anything."* ***CPSW 4*** *"A percentage wouldn't be helpful in court, because it would actually lead to more uncertainty. There would always be an argument to have which would detract away from the purpose."* ***Judge 4*** *"I appreciate what it is, there could be a range but what would low, medium, high then be? I suppose that's the problem."* ***Police Officer 7***
Confidence intervals	*"I would like to know the variance, that's what we would like to know, that's what we would need to know…"* ***CPSW 3*** *"I do like the number as a percentage but I do like to know the confidence interval as well…"* ***Clinician 7*** *"I think it might be a bit too much information."* ***CPSW 6***
Additional suggestions	*"There should be a little caveat statement there saying that it can go both ways, the higher it is the more likely it is to be, but a low one doesn't exclude it."* ***Clinician 23*** *"I think you need some sort of disclaimer on it about this needs to be used as part of a full assessment."* ***Clinician 8*** *"You want to know where it's come from, what's the research basis behind it, what's the evidence behind it, how much can you trust it. Now that could be a short blurb and then links to the publications, the literature that supports this."* ***Clinician 2*** *"I would want to know the data behind it because it is obviously chunking and splitting the data in different ways, maybe if you have got all the individual data, you could list everyone who fell outside the non-accidental injury bracket?"* ***Clinician 24*** *"If you can say here's a big block of how many of these kids were deliberately injured compared to a little smidge of kids, you could almost support it perhaps with a quick graphic to go big block is battered kids, small block is unfortunate accident."* ***Police Officer 4***

Figure 1. Flowchart of child protection professionals participating in an evaluation of the acceptability of the Predicting Abusive Head Trauma clinical prediction tool



CPSWs = Child Protection Social Workers

*In the United Kingdom, community pediatricians are doctors with expertise in working with vulnerable groups of children and their carers, including those who are at risk of abuse or are being abused. They have additional training in safeguarding.

Appendix 1.

Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist

Developed from:

Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*. 2007. Volume 19, Number 6: pp. 349 – 357

Item number	Guide questions/description	Reported in
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer/facilitator	Which author/s conducted the interview or focus group?	Laura Cowley Methods – Data Collection
2. Credentials	What were the researcher's credentials? E.g. PhD, MD	MSc Neuropsychology BSc (Hons) Psychology
3. Occupation	What was their occupation at the time of study?	PhD student Methods – Data Collection
4. Gender	Was the researcher male or female?	Female
5. Experience and training	What experience or training did the researcher have?	The researcher received substantial experience with qualitative research methods in her undergraduate and postgraduate degrees, and undertook a number of qualitative research projects as part of these. This experience was supplemented with the following recent training courses: "Interviewing in Social Science Research" (2015), "Qualitative Analysis Software" (2015), "Qualitative Data Analysis" (2016) and "Interpreting and writing up your Qualitative findings" (2016) Methods – Data Collection
Relationship with participants		
6. Relationship established	Was a relationship established prior to study commencement?	No Methods – Data Collection

7. Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	Participants were informed that the research study was being conducted as part of the researcher's PhD project via the Information Sheet
8. Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. bias, assumptions, reasons and interests in the research topic	The interviewer is a PhD student researching abusive head trauma and considered how her assumptions may influence the interviews and findings Methods – Data Collection
Domain 2: study design		
Theoretical framework		
9. Methodological orientation and Theory	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	Thematic analysis based on a general inductive approach Methods – Data Analysis
Participant selection		
10. Sampling	How were participants selected? e.g. purposive, convenience, consecutive, snowball	Purposive and snowball sampling to identify professionals involved in suspected AHT cases Methods – Participant recruitment
11. Method of approach	How were participants approached? e.g. face-to-face, telephone, mail, email	Email, or letters to judges Methods – Participant recruitment
12. Sample size	How many participants were in the study?	56 Table 1 and Figure 1
13. Non-participation	How many people refused to participate or dropped out? Reasons?	97 invited 76 registered 56 took part Figure 1
Setting		
14. Setting of data collection	Where was the data collected? E.g. home, clinic, workplace	Participants' workplace Methods – Data Collection
15. Presence of non-participants	Was anyone else present besides the participants and researchers?	Yes MF to record field notes Methods – Data Collection
16. Description of sample	What are the important characteristics of the	Table 1

	sample? e.g. demographic data, date	
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	The schedule included open-ended questions, prompts and clarifying questions and was piloted with two people Methods – Interview Schedule Development
18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	No Methods – Data Collection
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	Audio recording Methods – Data Collection
20. Field notes	Were field notes made during and/or after the interview or focus group?	Yes Methods – Data Collection
21. Duration	What was the duration of the interview or focus group?	45 minutes Methods – Data Collection
22. Data saturation	Was data saturation discussed?	Yes data saturation was verified using the constant comparative method Methods – Data Analysis
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	Only for two people who declined to be audio recorded Methods – Data Collection
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	Three Methods – Data Analysis
25. Description of the coding tree	Did authors provide a description of the coding tree?	The analytic framework is provided in Appendix 3
26. Derivation of themes	Were themes identified in advance or derived from the data?	Derived inductively from the data Methods – Data Analysis
27. Software	What software, if applicable, was used to manage the data?	NVivo 10 Methods – Data Analysis
28. Participant checking	Did participants provide feedback on the findings?	No Methods – Data Analysis
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes/findings? Was each	Quotations were presented and each participant was identified according to their

	quotation identified? e.g. participant number	professional group and participant number Results
30. Data and findings consistent	Was there consistency between the data presented and the findings?	The use of the constant comparative method ensured that quotations under each theme and category were reviewed for consistency and coherence Results
31. Clarity of major themes	Were major themes clearly presented in the findings?	All themes and categories identified during data analysis were presented in the results Results
32. Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Yes, discrepant cases and minor themes are discussed throughout the results Results

Appendix 2. Semi-structured Interview Schedule

Introduction

Hello my name is Laura and I will be interviewing you today. Thank you for being willing to take part in this project. Firstly, I would like to ask you for permission to audio record this interview. The main reasons for this are to ensure that the data collected is detailed and accurate and to facilitate data analysis. I would like to assure you that everything you say will remain completely confidential and only the immediate study team will have access to the audio recording and transcript. I am going to be showing you a tool that the research team have developed to estimate the probability of abuse in head-injured children, and ask you some questions about your thoughts on the usefulness of this tool. Do you have any questions before we proceed?

Explanation of the PredAHT clinical prediction tool

We have developed a clinical tool to estimate the probability of abusive head trauma in young children with head injuries, based on varying combinations of six clinical features. Each of the six features were included in the tool based on evidence from a systematic review of the literature, and were assigned a different weighting relative to their significance in a statistical model. The tool is intended for clinicians to complete, and we think it may be useful if they could communicate the results to other professionals who are involved in the child protection process, such as pathologists, police officers, child protection social workers, and legal professionals. It is intended for consideration alongside everything else that is known about each case, and should not be used as a diagnostic tool.

Participants' prior knowledge of clinical prediction tools

- Are you familiar with clinical prediction/decision rules?
- **If yes:** What is your opinion of them in general?

Evaluations of PredAHT

Do you think PredAHT would be useful for your investigations/practise/decision making?

- Why/why not?
- Could you tell me specifically how it would be useful for you?
- Perhaps you haven't had much/any experience with suspected abusive head trauma cases before?
- Perhaps it would give you more confidence in your decisions?

Can you think of a recent case in which PredAHT would have been useful to you?

- In what way would it have been useful to you?

Can you think of any factors that would make you more or less likely to use PredAHT?

- Perhaps if you knew how it was developed?

- Perhaps if it was supported by your colleagues?
- Perhaps if you were able to use it alongside other evidence?

Do you think there would be any risks involved in using PredAHT?

- What risks?

Do you think PredAHT would assist you in your discussions with other professionals involved in a case?

- How?/Why not?
- Perhaps it would be useful in a strategy meeting?

For clinicians only

When in the process of your investigations would PredAHT help you?

- At first presentation?
- Do you think it would be useful if you had missing data?
- Would it only be useful once all of the information about the features included in PredAHT was known?
- Would you use PredAHT to direct further examinations for example a skeletal survey?
- If further information became known regarding the features included in PredAHT, would you use it a second time to assess the change in the score?

Who do you think should complete PredAHT?

How do you think PredAHT could be integrated into existing hospital systems?

Presentation of the calculated probabilities

What do you think about the score being presented in terms of a percentage probability?

- Is there any other way you would want the likelihood of abuse expressed? Why/Why not?
- Would you prefer it to be translated into a low, medium or high likelihood of abuse?
- Would you want to know the estimate of uncertainty around the score (confidence interval)?
- Do you have any other suggestions about how the results should be presented?

Interpretation of probabilities in the context of suspected AHT

- How great would the likelihood of abuse have to be in percentage terms for you to take further action?
- What would a low, medium or high likelihood of abuse mean to you in percentage terms?
- What does the phrase 'on the balance of probabilities' mean to you in percentage terms?

Closure

We seem to have covered a great deal of ground and you have been very patient. However do you think there is anything that we have missed out that might be relevant or important? Do you have any other comments about what we've discussed or about the research as a whole?

Appendix 3. Analytic Framework

Theme	Category	Subcategories and Definitions
Evaluations of PredAHT	**Potential benefits of PredAHT**	**Objectivity:** any perceptions of PredAHT as being free of personal biases or evidence-based, and the advantages of this for investigating suspected AHT cases A**wareness:** any comments regarding PredAHT as useful for heightening awareness of the possibility of AHT (or nAHT), or reinforcing, increasing (or decreasing) concerns or suspicions about possible AHT **Reassurance:** comments about how PredAHT could provide assurance or confidence that participants' concerns, suspicions or investigations (or lack thereof) are justified; accounts of how PredAHT may be useful to back up or support participants' professional opinions or judgment **Rationalization of decisions:** any comments regarding PredAHT as useful for helping participants to explain, justify or rationalize their decision-making in suspected AHT cases **Standardization of clinical investigation:** any comments regarding PredAHT as useful for prompting clinicians to perform a clinical work-up to look for fractures or retinal hemorrhages, or modifying the clinical investigation e.g. by double-checking results **Justification for further action:** any comments about PredAHT as useful for justifying further action, investigations or assessments, including clinical/social work investigations, requests for charging decisions, or additional resources **Contributing to 'the bigger picture':** discussions about PredAHT as an additional factor to be considered as part of the wider picture; comments about PredAHT being useful for piecing parts of the clinical information together **Communication:** discussions about PredAHT as useful for facilitating communication between professionals; comments about whether the participants would share the result of PredAHT with their colleagues; any references to how the scores might be discussed at multi-agency strategy meetings or as part of information sharing **Training:** discussions about the benefits of being aware of the six clinical features included in PredAHT as potential indicators for AHT, or of PredAHT being useful for peer review or training purposes **Useful for 'grey' cases:** discussions regarding how PredAHT may be beneficial when working on 'grey' cases, where there is considerable uncertainty surrounding the diagnosis

		Useful for the less experienced: any comments regarding the potential benefits of PredAHT for those who have had little experience working in the child protection arena
	Potential risks of PredAHT	**Over-reliance:** any concerns that professionals may place too much reliance on PredAHT to aid their decision-making in suspected AHT cases **False reassurance from a low score:** any remarks about a 'low score' e.g. 14% instilling a false sense of security; concerns that appropriate investigations would not be carried out if a low score was obtained **May not be used as intended:** any concerns that the tool would be improperly used; concerns that it may be used in isolation, without consideration of other factors relevant to AHT cases such as caregiver provided history or social history; concerns that the tool is too reductionist or crude **Accuracy of PredAHT:** any comments relating to the accuracy, sensitivity or specificity of PredAHT, discussions about false positives or false negatives and related implications **Irrelevant:** any comments about PredAHT being irrelevant or not particularly useful for participants' decision-making, remarks that it may not add much to what is already known **Features not included in PredAHT:** any important features that the participants' feel are missing and why e.g. skull fractures **Age, number, location, pattern and severity of injuries:** discussions regarding any information or details about suspected AHT cases that cannot be taken into account by PredAHT and the impact this may have on decision-making; comments relating to the inability of PredAHT to account for or distinguish between repeated or multiple injuries that may have been sustained over time, e.g. healing fractures; the precise location or pattern of the injuries e.g. posterior rib fracture; the seriousness of the injuries **Introduces bias:** remarks that PredAHT may condition decision-making or introduce apparent or unconscious bias
	Provisos for the use of PredAHT	**If accepted by colleagues:** comments regarding the acceptability of PredAHT to colleagues in the medical profession or colleagues in their own or other agencies as a stipulation for use **Alongside professional judgement/other factors:** any remarks regarding PredAHT being an additional piece of information to make use of in conjunction with other factors relating to the case as well as participants' professional opinion **If kept up to date:** comments about the need for PredAHT to be regularly updated in light of current evidence

		Definition of the features: participants' understandings of the features included in PredAHT; remarks about the need for the features to be explicitly defined **Understanding how PredAHT works:** remarks about the desire to understand how PredAHT was developed, how it works and how it should be used; comments about training requirements **Quality of the data:** comments about the need to appraise the quality of the underlying data used to derive PredAHT **Agreeing accepted risk thresholds:** discussions about differing risk thresholds and the need for a consistent interpretation of the scores between colleagues and between agencies
	Practical use of PredAHT	**Usability/simplicity:** any comments about the ease of use of PredAHT, any potential barriers to completing it e.g. time/complexity **Hospital settings it would be useful:** comments about the settings in which PredAHT could be used e.g. district hospitals, the paediatric intensive care unit, the emergency department **Stages of the assessment process:** comments about the relative usefulness of PredAHT at the different stages of the investigative/assessment process; remarks about the best time point to use it **Who should complete it:** discussions about the best person to be responsible for completing PredAHT e.g. admitting consultant/safeguarding professional **Integration into the clinical workflow:** any remarks about how PredAHT might be implemented into existing hospital systems; comments about the most appropriate medium by which to use it e.g. computer/phone **Prior probability:** comments about estimating a prior probability of AHT and whether this would be difficult to do in practice; remarks about needing more information or guidance to complete this aspect of the tool; comments about the impact of incorporating a prior probability of AHT into the tool
	Use of PredAHT in court	**Evidence-based:** comments about PredAHT being useful in a court setting because it is based on evidence or has been validated **Standards of proof:** discussions about the value of PredAHT in court in relation to the standards of proof adhered to in different court settings; concern that a high score may equate to the term "beyond all reasonable doubt" and act as a deciding factor in a conviction **Restrictive:** comments that PredAHT could not account for every factor in every case and so would be dismissed; comments that PredAHT can only take into account a limited number of clinical features

		Cross-examination: any remarks about counter-arguments that may arise as a result of using the tool in court e.g. criticisms of how the tool was developed, claims that the case falls into the reverse or 'flip' of the probability given by the tool; the need for those presenting the results to understand and explain how the tool was constructed **Identifying the perpetrator:** comments about the inability of PredAHT to identify a possible perpetrator in suspected AHT cases **Historical child protection court cases involving statistical evidence:** any discussions about previous court cases in child protection that have involved the use of quantitative tools or probability theory, and the impact of these cases on participants' thoughts about using PredAHT in court **Medical court report:** remarks about how the score could be useful as part of the wider medical report submitted to the courts
Presentation of calculated probabilities	**Percentage probabilities versus broad risk categories:**	Any comments regarding whether the results should be expressed as a precise numerical score, or using broad categories e.g. 'low, medium and high likelihood of abuse'; any explanations for participants' preferred choices
	Confidence intervals	Any discussions about participants' understanding of confidence intervals; any comments about whether or not they should be provided alongside the score and why
	Additional suggestions	Any other suggestions for expressing the results generated by PredAHT, including associated disclaimers, background information or visual aids
Interpretations of probabilities in the context of suspected AHT	**Threshold criteria**	Any comments about participants' accepted probability thresholds for abuse and non-abuse; discussions about thresholds for continuing or stopping investigations/assessments; perceptions of the percentage equivalents and meanings of different risk categories such as 'likely' and 'unlikely'
	Comments about scores generated by PredAHT	Comments about the percentages that PredAHT gives for different combinations of features; participants' opinions of the scores associated with any cases they are describing